

# Lecture 7: Further Algorithms and Theory

Simo Särkkä

Department of Biomedical Engineering and Computational Science  
Helsinki University of Technology

May 5, 2009

- 1 Further Filtering Algorithms
- 2 Continuous-Discrete-Time EKF
- 3 General Continuous-Discrete-Time Filtering
- 4 Continuous-Time Filtering
- 5 Linear Stochastic Differential Equations
- 6 What is Beyond This?
- 7 Summary

# Multiple Model Kalman filtering

- Algorithm for estimating **true model or its parameter** from a **finite set of alternatives**.
- Assume that we are given  **$N$  possible dynamic/measurement models**, and **one of them is true**.
- If  **$s$**  is the **model index**, the problem can be written in form:

$$\begin{aligned}s &\sim P(s) \\ \mathbf{x}_k &= \mathbf{A}(s) \mathbf{x}_{k-1} + \mathbf{q}_{k-1} \\ \mathbf{y}_k &= \mathbf{H}(s) \mathbf{x}_k + \mathbf{r}_k,\end{aligned}$$

where  $\mathbf{q}_{k-1} \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}(s))$  and  $\mathbf{r}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}(s))$ .

- Can be solved in closed form with  **$s$  parallel Kalman filters**.

# Interacting Multiple Models (IMM) Filter

- Assume that we have  $N$  possible models, but the true model is assumed to change in time.
- If the model index  $s_k$  is modeled as Markov chain, we have:

$$s_k \sim P(s_k | s_{k-1})$$

$$\mathbf{x}_k = \mathbf{A}(s_{k-1}) \mathbf{x}_{k-1} + \mathbf{q}_{k-1}$$

$$\mathbf{y}_k = \mathbf{H}(s_k) \mathbf{x}_k + \mathbf{r}_k,$$

- Closed form solution would require running Kalman filters for each possible history  $s_{1:k} \Rightarrow N^k$  filters, not feasible.
- Interacting Multiple Models (IMM) filter is an approximation algorithm, which uses only  $N$  parallel Kalman filters.

# Variational Kalman Smoother

- **Variation Bayesian analysis** based framework for estimating the **parameters of linear state space models**.
- Idea: Fix  $\mathbf{Q} = \mathbf{I}$  and assume that the joint distribution of states  $\mathbf{x}_1, \dots, \mathbf{x}_T$  and parameters  $\mathbf{A}, \mathbf{H}, \mathbf{R}$  is **approximately separable**:

$$\begin{aligned} p(\mathbf{x}_1, \dots, \mathbf{x}_T, \mathbf{A}, \mathbf{H}, \mathbf{R} \mid \mathbf{y}_1, \dots, \mathbf{y}_T) \\ \approx p(\mathbf{x}_1, \dots, \mathbf{x}_T \mid \mathbf{y}_1, \dots, \mathbf{y}_T) p(\mathbf{A}, \mathbf{H}, \mathbf{R} \mid \mathbf{y}_1, \dots, \mathbf{y}_T). \end{aligned}$$

- The resulting **EM-algorithm** consist of alternating steps of smoothing with fixed parameters and estimation of new parameter values.
- The general equations of the algorithm are quite complicated and assume that all the model parameters are to be estimated.

# Recursive Variational Bayesian Estimation of Noise Variances

- Algorithm for estimating unknown **time-varying measurement variances**.
- Assume that the joint filtering distribution of state and measurement noise variance is **approximately separable**:

$$p(\mathbf{x}_k, \sigma_k^2 | y_1, \dots, y_k) \approx p(\mathbf{x}_k | y_1, \dots, y_k) p(\sigma_k^2 | y_1, \dots, y_k)$$

- Variational Bayesian analysis** leads to algorithm, where the natural representation is

$$p(\sigma_k^2 | y_1, \dots, y_k) = \text{InvGamma}(\sigma_k^2 | \alpha_k, \beta_k)$$

$$p(\mathbf{x}_k | y_1, \dots, y_k) = \mathbf{N}(\mathbf{x}_k | \mathbf{m}_k, \mathbf{P}_k).$$

- The update step consists of a **fixed-point iteration** for computing new  $\alpha_k, \beta_k, \mathbf{m}_k, \mathbf{P}_k$  from the old ones.

- Outlier Rejection / Clutter Modeling:
  - Probabilistic Data Association (PDA)
  - Monte Carlo Data Association (MCDA)
  - Multiple hypothesis tracking (MHT)
- Multiple Target Tracking
  - Multiple hypothesis tracking (MHT)
  - Joint Probabilistic Data Association (JPDA)
  - Rao-Blackwellized Particle Filtering (RBMCPDA) for Multiple Target Tracking

# Continuous-Discrete Pendulum Model

- Consider the **pendulum model**, which was first stated as

$$\begin{aligned}d^2\theta/dt^2 &= -g \sin(\theta) + w(t) \\ y_k &= \sin(\theta(t_k)) + r_k,\end{aligned}$$

where  $w(t)$  is "Gaussian white noise" and  $r_k \sim N(0, \sigma^2)$ .

- With state  $\mathbf{x} = (\theta, d\theta/dt)$ , the model is of the **abstract form**

$$\begin{aligned}d\mathbf{x}/dt &= \mathbf{f}(\mathbf{x}) + \mathbf{w}(t) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}(t_k)) + \mathbf{r}_k\end{aligned}$$

where  $\mathbf{w}(t)$  has the covariance (spectral density)  $\mathbf{Q}_c$ .

- Continuous-time dynamics + discrete-time measurement = **Continuous-discrete (-time) filtering model**.

- Previously we assumed that the **measurements are obtained at times**  $t_k = 0, \Delta t, 2\Delta t, \dots$
- The state space model was then **Euler-discretized** as

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{f}(\mathbf{x}_{k-1}) \Delta t + \mathbf{q}_{k-1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k$$

- But what should be the **variance of  $\mathbf{q}_k$** ?
- **Consistency**: The same variance for single step of length  $\Delta t$ , and 2 steps of length  $\Delta t/2$ :

$$\mathbf{q}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_c \Delta t)$$

- Now the **Extended Kalman filter (EKF)** for this model is
  - Prediction:

$$\begin{aligned}\mathbf{m}_k^- &= \mathbf{m}_{k-1} + \mathbf{f}(\mathbf{m}_{k-1}) \Delta t \\ \mathbf{P}_k^- &= (\mathbf{I} + \mathbf{F} \Delta t) \mathbf{P}_{k-1} (\mathbf{I} + \mathbf{F} \Delta t)^T + \mathbf{Q}_c \Delta t \\ &= \mathbf{P}_{k-1} + \mathbf{F} \mathbf{P}_{k-1} \Delta t + \mathbf{P}_{k-1} \mathbf{F}^T \Delta t \\ &\quad + \mathbf{F} \mathbf{P}_{k-1} \mathbf{F}^T \Delta t^2 + \mathbf{Q}_c \Delta t\end{aligned}$$

- Update:

$$\begin{aligned}\mathbf{S}_k &= \mathbf{H}(\mathbf{m}_k^-) \mathbf{P}_k^- \mathbf{H}^T(\mathbf{m}_k^-) + \mathbf{R} \\ \mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}^T(\mathbf{m}_k^-) \mathbf{S}_k^{-1} \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k [\mathbf{y}_k - \mathbf{h}(\mathbf{m}_k^-)] \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T\end{aligned}$$

- But what happens if  $\Delta t$  is not “small”, that is, if we get measurements quite rarely?
  - We can use **more Euler steps** between measurements.
  - We can perform the **EKF prediction on each step**.
  - We can even compute the limit of **infinite number of steps**.
- If we let  $\delta t = \Delta t/n$ , the **prediction** becomes:

$$\hat{\mathbf{m}}_0 = \mathbf{m}_{k-1}; \quad \hat{\mathbf{P}}_0 = \mathbf{P}_{k-1}$$

for  $i = 1 \dots n$

$$\hat{\mathbf{m}}_i = \hat{\mathbf{m}}_{i-1} + \mathbf{f}(\hat{\mathbf{m}}_{i-1}) \delta t$$

$$\begin{aligned} \hat{\mathbf{P}}_i &= \hat{\mathbf{P}}_{i-1} + \mathbf{F} \hat{\mathbf{P}}_{i-1} \delta t + \hat{\mathbf{P}}_{i-1} \mathbf{F}^T \delta t \\ &\quad + \mathbf{F} \hat{\mathbf{P}}_{i-1} \mathbf{F}^T \delta t^2 + \mathbf{Q}_c \delta t \end{aligned}$$

end

$$\mathbf{m}_k^- = \hat{\mathbf{m}}_n; \quad \mathbf{P}_k^- = \hat{\mathbf{P}}_n.$$

- By **re-arranging the equations** in the for-loop, we get

$$(\hat{\mathbf{m}}_i - \hat{\mathbf{m}}_{i-1})/\delta t = \mathbf{f}(\hat{\mathbf{m}}_{i-1})$$

$$(\hat{\mathbf{P}}_i - \hat{\mathbf{P}}_{i-1})/\delta t = \mathbf{F} \hat{\mathbf{P}}_{i-1} + \hat{\mathbf{P}}_{i-1} \mathbf{F}^T + \mathbf{F} \hat{\mathbf{P}}_{i-1} \mathbf{F}^T \delta t + \mathbf{Q}_c$$

- In the limit  $\delta t \rightarrow 0$ , we get the **differential equations**

$$d\hat{\mathbf{m}}/dt = \mathbf{f}(\hat{\mathbf{m}}(t))$$

$$d\hat{\mathbf{P}}/dt = \mathbf{F}(\hat{\mathbf{m}}(t)) \hat{\mathbf{P}}(t) + \hat{\mathbf{P}}(t) \mathbf{F}^T(\hat{\mathbf{m}}(t)) + \mathbf{Q}_c$$

- The **initial conditions** are

$$\hat{\mathbf{m}}(0) = \mathbf{m}_{k-1}$$

$$\hat{\mathbf{P}}(0) = \mathbf{P}_{k-1}$$

- The **final prediction** is

$$\mathbf{m}_k^- = \hat{\mathbf{m}}(\Delta t)$$

$$\mathbf{P}_k^- = \hat{\mathbf{P}}(\Delta t)$$

## Continuous-Discrete EKF

- **Prediction:** between the measurements integrate the following differential equations from  $t_{k-1}$  to  $t_k$ :

$$d\mathbf{m}/dt = \mathbf{f}(\mathbf{m}(t))$$

$$d\mathbf{P}/dt = \mathbf{F}(\mathbf{m}(t)) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{F}^T(\mathbf{m}(t)) + \mathbf{Q}_c$$

- **Update:** at measurements do the EKF update

$$\mathbf{S}_k = \mathbf{H}(\mathbf{m}_k^-) \mathbf{P}_k^- \mathbf{H}^T(\mathbf{m}_k^-) + \mathbf{R}$$

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T(\mathbf{m}_k^-) \mathbf{S}_k^{-1}$$

$$\mathbf{m}_k = \mathbf{m}_k^- + \mathbf{K}_k [\mathbf{y}_k - \mathbf{h}(\mathbf{m}_k^-)]$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T,$$

where  $\mathbf{m}_k^-$  and  $\mathbf{P}_k^-$  are the results of the prediction step.

- The equations

$$d\mathbf{m}/dt = \mathbf{f}(\mathbf{m}(t))$$

$$d\mathbf{P}/dt = \mathbf{F}(\mathbf{m}(t))\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(\mathbf{m}(t)) + \mathbf{Q}_c$$

actually generate a **Gaussian process approximation**  $\mathbf{x}(t) \sim \mathbf{N}(\mathbf{m}(t), \mathbf{P}(t))$  to the solution of **non-linear stochastic differential equation (SDE)**

$$d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}) + \mathbf{w}(t)$$

- We could also use **statistical linearization or unscented transform** and get a bit different limiting differential equations.
- Also possible to generate **particle approximations** by a continuous-time version of importance sampling (based on Girsanov theorem).

# More general SDE Theory

- The **most general SDE model** usually considered is of the form

$$d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}) + \mathbf{L}(\mathbf{x}) \mathbf{w}(t)$$

- Formally,  $\mathbf{w}(t)$  is a **Gaussian white noise process** with zero mean and covariance function

$$E[\mathbf{w}(t) \mathbf{w}^T(t')] = \mathbf{Q}_c \delta(t' - t)$$

- The distribution  $p(\mathbf{x}(t))$  is **non-Gaussian** and it is given by the following **partial differential equation**:

$$\frac{\partial p}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} (f_i(\mathbf{x}) p) + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} ([\mathbf{L} \mathbf{Q} \mathbf{L}^T]_{ij} p)$$

- Known as **Fokker-Planck** equation or **Kolmogorov forward** equation.

# More general SDE Theory (cont.)

- In more **rigorous theory**, we actually must interpret the SDE as **integral equation**

$$\mathbf{x}(t) - \mathbf{x}(s) = \int_s^t \mathbf{f}(\mathbf{x}) dt + \int_s^t \mathbf{L}(\mathbf{x}) \mathbf{w}(t) dt$$

- In **Ito's** theory of SDE's the second integral is defined as **stochastic integral** w.r.t. **Brownian motion**  $\beta(t)$ :

$$\mathbf{x}(t) - \mathbf{x}(s) = \int_s^t \mathbf{f}(\mathbf{x}) dt + \int_s^t \mathbf{L}(\mathbf{x}) d\beta(t)$$

i.e., formally  $\mathbf{w}(t) dt = d\beta(t)$  or  $\mathbf{w}(t) = d\beta(t)/dt$

- However, **Brownian motion is nowhere differentiable!**
- Brownian motion is also called as **Wiener process**.

- In **stochastics**, the integral equation is often written as

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}) dt + \mathbf{L}(\mathbf{x}) d\beta(t)$$

- In **engineering** (control theory, physics) it is customary to formally divide with  $dt$  to get

$$d\mathbf{x}(t)/dt = \mathbf{f}(\mathbf{x}) + \mathbf{L}(\mathbf{x}) \mathbf{w}(t)$$

- So called **Stratonovich's theory** is more consistent with this **white noise interpretation** than Ito's theory.
- In mathematical sense Stratonovich's theory defines the **stochastic integral**  $\int_s^t \mathbf{L}(\mathbf{x}) d\beta(t)$  a bit differently – also the Fokker-Planck equation is different.

# Cautions About White Noise

- White noise is actually **only formally** defined as **derivative of Brownian motion**.
- White noise can only be defined in **distributional sense** – for this reason **non-linear functions** of it  $g(\mathbf{w}(t))$  are **not well-defined**.
- For this reason, the following more general type of SDE **does not make sense**:

$$d\mathbf{x}(t)/dt = \mathbf{f}(\mathbf{x}, \mathbf{w})$$

- We must also be **careful** in interpreting the **multiplicative term** in the equation

$$d\mathbf{x}(t)/dt = \mathbf{f}(\mathbf{x}) + \mathbf{L}(\mathbf{x}) \mathbf{w}(t)$$

## Optimal continuous-discrete filter

- 1 **Prediction step:** Solve the Kolmogorov-forward (Fokker-Planck) partial differential equation.

$$\frac{\partial p}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} (f_i(\mathbf{x}) p) + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} ([\mathbf{L} \mathbf{Q} \mathbf{L}^T]_{ij} p)$$

- 2 **Update step:** Apply the Bayes' rule.

$$p(\mathbf{x}(t_k) | \mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k | \mathbf{x}(t_k)) p(\mathbf{x}(t_k) | \mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k | \mathbf{x}(t_k)) p(\mathbf{x}(t_k) | \mathbf{y}_{1:k-1}) d\mathbf{x}(t_k)}$$

# General Continuous-Time Filtering

- We could also model the **measurements** as a **continuous-time process**:

$$\begin{aligned}d\mathbf{x}/dt &= \mathbf{f}(\mathbf{x}) + \mathbf{L}(\mathbf{x}) \mathbf{w}(t) \\ \mathbf{y} &= \mathbf{h}(\mathbf{x}) + \mathbf{n}(t)\end{aligned}$$

- Again, one must be very careful in **interpreting the white noise** processes  $\mathbf{w}(t)$  and  $\mathbf{n}(t)$ .
- The filtering equations become a **stochastic partial differential equation** (SPDE) called **Kushner-Stratonovich equation**.
- The equation for the unnormalized filtering density is called the **Zakai equation**, which also is a SPDE.
- It is also possible to take the **continuous-time limit of the Bayesian smoothing equations** (result is a PDE).

- If the system is **linear**

$$d\mathbf{x}/dt = \mathbf{F}\mathbf{x} + \mathbf{w}(t)$$

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}(t)$$

we get the **continuous-time Kalman-Bucy filter**:

$$d\mathbf{m}/dt = \mathbf{F}\mathbf{m} + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{m})$$

$$d\mathbf{P}/dt = \mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^T + \mathbf{Q}_c - \mathbf{K}\mathbf{R}\mathbf{K}^T,$$

where  $\mathbf{K} = \mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}$ .

- The **stationary solution** to these equations is equivalent to the continuous-time **Wiener filter**.
- **Non-linear extensions** (EKF, SLF, UKF, etc.) can be obtained similarly to the discrete-time case.

- Let's return to **linear stochastic differential equations**:

$$d\mathbf{x}/dt = \mathbf{F}\mathbf{x} + \mathbf{w}$$

- Assume that  $\mathbf{F}$  is **time-independent**. For example, in car-tracking model we had a model of this type.
- Given  $\mathbf{x}(0)$  we can now actually **solve the equation**

$$\mathbf{x}(t) = \exp(t\mathbf{F})\mathbf{x}(0) + \int_0^t \exp((t-s)\mathbf{F})\mathbf{w}(s) ds,$$

where  $\exp(\cdot)$  is the matrix exponential function:

$$\exp(t\mathbf{F}) = \mathbf{I} + t\mathbf{F} + \frac{1}{2!}t^2\mathbf{F}^2 + \frac{1}{3!}t^3\mathbf{F}^3 + \dots$$

- Note that we are treating  $\mathbf{w}(s)$  **as an ordinary function**, which is not generally justified!

- We can also **solve the equation on predefined time points**  $t_1, t_2, \dots$  as follows:

$$\mathbf{x}(t_k) = \exp((t_k - t_{k-1}) \mathbf{F}) \mathbf{x}(t_{k-1}) + \int_{t_{k-1}}^{t_k} \exp((t_k - s) \mathbf{F}) \mathbf{w}(s) ds$$

- The **first term** is of the form  $\mathbf{A} \mathbf{x}(t_{k-1})$ , where the matrix is a known constant  $\mathbf{A} = \exp(\Delta t \mathbf{F})$ .
- The **second term** is a zero mean Gaussian random variable and its covariance can be calculated as:

$$\begin{aligned} \mathbf{Q} &= \int_{t_{k-1}}^{t_k} \exp((t_k - s) \mathbf{F}) \mathbf{Q}_c \exp((t_k - s) \mathbf{F})^T ds \\ &= \int_0^{\Delta t} \exp((\Delta t - s) \mathbf{F}) \mathbf{Q}_c \exp((\Delta t - s) \mathbf{F})^T ds \end{aligned}$$

- Thus the **continuous-time system** is in a sense **equivalent to the discrete-time system**

$$\mathbf{x}(t_k) = \mathbf{A} \mathbf{x}(t_{k-1}) + \mathbf{q}_k$$

where  $\mathbf{q}_k \sim N(\mathbf{0}, \mathbf{Q})$  and

$$\mathbf{A} = \exp(\Delta t \mathbf{A})$$

$$\mathbf{Q} = \int_0^{\Delta t} \exp((\Delta t - s) \mathbf{F}) \mathbf{Q}_c \exp((\Delta t - s) \mathbf{F})^T ds$$

- An analogous **equivalent discretization** is also possible with **time-varying linear stochastic differential equation models**.
- A **continuous-discrete Kalman filter** can be always **implemented as a discrete-time Kalman filter** by forming the equivalent discrete-time system.

# Wiener Velocity Model

- For example, consider the **Wiener velocity model** (= white noise acceleration model):

$$d^2x/dt^2 = w(t),$$

which is equivalent to the **state space model**

$$d\mathbf{x}/dt = \mathbf{F} \mathbf{x} + \mathbf{w}$$

with  $\mathbf{F} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $\mathbf{x} = (x, dx/dt)$ ,  $\mathbf{Q}_c = \begin{pmatrix} 0 & 0 \\ 0 & q \end{pmatrix}$ .

- Then we have

$$\mathbf{A} = \exp(\Delta t \mathbf{F}) = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{Q} &= \int_0^{\Delta t} \exp((\Delta t - s) \mathbf{F}) \mathbf{Q}_c \exp((\Delta t - s) \mathbf{F})^T ds \\ &= \begin{pmatrix} \Delta t^3/3 q & \Delta t^2/2 q \\ \Delta t^2/2 q & \Delta t q \end{pmatrix} \end{aligned}$$

which might look familiar.

# Mean and Covariance Differential Equations

- Note that in the **linear (time-invariant) case**

$$dx/dt = \mathbf{F} \mathbf{x} + \mathbf{w}$$

we could also write down the **differential equations**

$$d\mathbf{m}/dt = \mathbf{F} \mathbf{m}$$

$$d\mathbf{P}/dt = \mathbf{F} \mathbf{P} + \mathbf{P} \mathbf{F}^T + \mathbf{Q}_c$$

which exactly give the **evolution of mean and covariance**.

- The **solutions** of these equations are

$$\mathbf{m}(t) = \exp(t \mathbf{F}) \mathbf{m}_0$$

$$\mathbf{P}(t) = \exp(t \mathbf{F}) \mathbf{P}_0 \exp(t \mathbf{F})^T$$

$$+ \int_0^t \exp((t-s) \mathbf{F}) \mathbf{Q}_c \exp((t-s) \mathbf{F})^T ds,$$

which are consistent with the previous results.

- Assume that the physical system can be modeled with **differential equation with input  $\mathbf{u}$**

$$d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

- Determine  $\mathbf{u}(t)$  such that  $\mathbf{x}(t)$  and  $\mathbf{u}(t)$  **satisfy certain constraints and minimize a cost functional**.
- For example, **steer a space craft to moon** such that the **consumed of fuel is minimized**.
- If the system is **linear and cost function quadratic**, we get **linear quadratic controller** (or regulator).

- Assume that the **system model is stochastic**:

$$d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, \mathbf{u}) + \mathbf{w}(t)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k)) + \mathbf{r}_k$$

- **Given only the measurements  $\mathbf{y}_k$ , find  $\mathbf{u}(t)$  such that  $\mathbf{x}(t)$  and  $\mathbf{u}(t)$  satisfy the constraints and minimize a cost function.**
- If linear Gaussian, we have  
Linear Quadratic (LQ) controller + Kalman filter = Linear Quadratic Gaussian (LQG) controller
- In general, **not** simply a combination of **optimal filter and deterministic optimal controller**.
- **Model Predictive Control (MPC)** is a well-known approximation algorithm for constrained problems.

- Infinite dimensional generalization of state space model is the stochastic partial differential equation (SPDE)

$$\frac{\partial \mathbf{x}(t, \mathbf{r})}{\partial t} = \mathcal{F}_r \mathbf{x}(t, \mathbf{r}) + \mathcal{L}_r \mathbf{w}(t, \mathbf{r}),$$

where  $\mathcal{F}_r$  and  $\mathcal{L}_r$  are linear operators (e.g. integro-differential operators) in  $\mathbf{r}$ -variable and  $\mathbf{w}(\cdot)$  is a time-space white noise.

- Practically every SPDE can be converted into this form with respect to any variable (which is relabeled as  $t$ ).
- For example, stochastic heat equation

$$\frac{\partial x(t, r)}{\partial t} = \frac{\partial^2 x(t, r)}{\partial r^2} + w(t, r).$$

# Spatially Distributed Systems (cont.)

- The **solution to the SPDE** is analogous to finite-dimensional case:

$$\mathbf{x}(t, \mathbf{r}) = \mathcal{U}_r(t) \mathbf{x}(0, \mathbf{r}) + \int_0^t \mathcal{U}_r(t-s) \mathcal{L}_r \mathbf{w}(s, \mathbf{r}) ds.$$

- $\mathcal{U}_r(t) = \exp(t \mathcal{F}_r)$  is the **evolution operator** – corresponds to **propagator** in quantum mechanics.
- **Spatio-temporal Gaussian process models** can be naturally formulated as linear SPDE's.
- Recursive Bayesian estimation with SPDE models lead to **infinite-dimensional Kalman filters and RTS smoothers**.
- SPDE's can be **approximated with finite models** by usage of **finite-differences or finite-element methods**.

# Summary

- Other filtering algorithms than EKF, SLF, UKF and PF are, for example, **multiple model Kalman filters and IMM algorithm**.
- **Specialized filtering algorithms** exist also, e.g., for parameter estimation, outlier rejection and multiple target tracking.
- **In continuous-discrete filtering**, the dynamic model is a continuous-time process and measurement are obtained at discrete times.
- **In continuous-discrete EKF, SLF and UKF** the continuous-time non-linear dynamic model is approximated as a **Gaussian process**.
- **In continuous-time filtering**, the both the dynamic and measurements models are continuous-time processes.
- The theories of continuous and continuous-discrete filtering are tied to the **theory of stochastic differential equations**.